DUF3380 domain from a *Salmonella* phage endolysin shows potent *N*-acetylmuramidase activity

Lorena Rodríguez-Rubio[1], Hans Gerstmans[1,2], Simon Thorpe[3], Stéphane Mesnage[4], Rob Lavigne[1]#, Yves Briers[2]#

[1]KU Leuven, Laboratory of Gene Technology, Kasteelpark Arenberg 21, box 2462, 3001 Leuven, Belgium

[2]Ghent University, Department of Applied Biosciences, Laboratory of Applied Biotechnology, Valentin Vaerwyckweg 1, 9000 Ghent, Belgium

[3]Faculty of Science, Mass Spectrometry Facility, University of Sheffield, Brook Hill Road, Sheffield S3 7HF, UK

[4]Krebs Institute, Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, UK

Running head: DUF3380 domain displays *N*-acetylmuramidase activity

Address correspondence to: Rob Lavigne, rob.lavigne@kuleuven.be, +32 (0)16 37 95 24 and Yves Briers, yves.briers@ugent.be, +32 (9) 243 24 53

ABSTRACT

Bacteriophage-encoded endolysins are highly diverse enzymes that cleave the bacterial peptidoglycan layer. Current research focuses on their potential applications in medicine, food conservation and as biotechnological tools. Despite the wealth of applications relying on the use of endolysin, little is known about the enzymatic properties of these enzymes, especially in case of endolysins of bacteriophages infecting Gram-negative species. Automated genome annotations therefore remain to be confirmed. Here, we report the biochemical analysis and cleavage site determination of a novel *Salmonella* bacteriophage endolysin, Gp110, which comprises an uncharacterized Domain of Unknown Function (DUF3380; pfam11860) domain in its C-terminus and shows the highest specific activity (34,240 U/µM) compared to fourteen previously characterized endolysins active against peptidoglycan from Gram-negative bacteria (corresponding to a 1.7- to 364-fold higher activity). Gp110 is a modular endolysin with an optimal pH of enzymatic activity at pH 8 and an elevated thermal resistance. Reversed phase-HPLC analysis coupled to mass spectrometry showed that DUF3380 has *N*-acetylmuramidase (lysozyme) activity cleaving the β-(1,4) glycosidic bond between *N*-acetylmuramic acid and *N*-acetylglucosamine residues. Gp110 is active against directly cross-linked peptidoglycan with various peptide stem compositions, making it an attractive enzyme to develop novel antimicrobial agents.

40 **IMPORTANCE**

41 We report the functional and biochemical characterization of the *Salmonella* phage endolysin
42 Gp110. This endolysin has a modular structure with an enzymatically active domain and a cell
43 wall binding domain. The enzymatic activity of this endolysin outstands all other endolysins
44 previously characterized using the same methods. A Domain of Unknown Function (DUF3380) is
45 responsible for this high enzymatic activity. We report that DUF3380 has a N-acetylmuramidase
46 activity against directly cross-linked peptidoglycan with various peptide stem compositions. This
47 experimentally verified activity will allow a better classification and understanding of endolysins
48 enzymatic activities which mostly are inferred by sequence similarities. Three-dimensional
49 structure predictions for Gp110 suggest a completely different fold compared to previous
50 enzymes with the same peptidoglycan cleavage specificity, making this endolysins quite unique.
51 All these features, combined with an increased thermal resistance, make Gp110 an attractive
52 candidate to engineer novel endolysin-based antibacterials.

53

INTRODUCTION

Endolysins are bacteriophage (phage)-encoded proteins synthesized at the end of the lytic infection cycle which degrade the peptidoglycan (PG) of the host bacterium to allow the viral progeny release (1). The specific activity and structure of these proteins have boosted their study as new antimicrobials against pathogens including multidrug resistant bacteria (2). Recently engineered fusions of an endolysin and a selected outer membrane permeabilizing peptide (Artilysin®s) were shown to display high activity against Gram-negative bacteria (3-5). In addition, the analysis of endolysins has also led to the development of new biotechnological tools for bacterial diagnostics and detection, among others (6).

Depending on their origin, the structure of endolysins varies. In general, most of the endolysins from phages infecting Gram-positive bacteria have a modular structure consisting of one or two N-terminal enzymatic active domains (EADs) and a C-terminal cell wall binding domain (CBD) separated by a short linker (7). In contrast, the vast majority of endolysins from phages infecting Gram-negative bacteria have a globular organization containing only an EAD, although a number of modular endolysins with different orientations of EADs and CBDs have been also described (4, 8). The CBDs are responsible for the recognition of the substrate and the high-affinity binding of these enzymes to the bacterial cell wall (9), whereas the EADs are responsible for the catalytic activity, i.e., the cleavage of specific bonds within the PG. The PG is a copolymer of alternating *N*-acetylmuramic acid (MurNAc) and *N*-acetylglucosamine (GlcNAc) residues linked by β-(1,4) glycosidic bonds. Lactyl groups of the MurNAc residues are substituted by a pentapeptide stem made of L- and D- amino acids which is highly conserved in

76 Gram-negative species (L-Ala-D-Glu-mDAP-D-Ala-D-Ala) but variable in Gram-positive species
77 (10). The *meso*-diaminopimelic acid (mDAP) is substituted by L-lysine in most Gram-positive
78 species, although it can still be found in *Bacillus* and *Listeria* and L-ornithine can be found in the
79 third position of the stem peptides in the PG of *Thermus thermophilus*, Spirochetes and
80 *Bifidobacterium globosum* (11).

81 Three groups of enzymatic activities have been associated with endolysins: i)
82 glycosidases, which include glucosaminidases (EC 3.2.1.52), muramidases or lysozymes (EC
83 3.2.1.17) and lytic transglycosylases (EC 4.2.2.n1) targeting the β-(1,4) glycosidic bonds of the
84 sugar backbone; ii) amidases (EC 3.5.1.28) which target the amide bond between the sugar
85 backbone and the peptide stems and iii) endopeptidases (EC 3.4.-.-) which hydrolyse the bond
86 between two amino acids. Only a limited number of studies have analyzed the PG bond cleaved
87 by endolysins (12-17) and most (automated) annotations of enzymatic specificity only rely on
88 sequence similarity. As a result, available databases contain inaccurate descriptions of
89 biochemical specificities. A major problem is associated with the fact that several endolysins
90 were referred to as lysozymes despite the lack of a biochemical characterization. A typical
91 example is the endolysin of the T7 bacteriophage originally named as "T7 lysozyme". This
92 erroneous designation still persists even though the T7 endolysin was experimentally
93 demonstrated to be a *N*-acetylmuramoyl-L-alanine amidase rather than *N*-acetylmuramidase
94 (or lysozyme) (18). Another typical example is the bacteriophage lambda lysozyme which has
95 been shown to display lytic transglycosidase activity (19). Furthermore, sequence similarity
96 used to assign a putative function to endolysins are sometimes very poor or limited, while Pfam

97 designations are often not updated. This can lead to discrepancies between *in silico* and
98 experimental results when the cleavage site determination is performed (20).

99 In this study, we report the functional and biochemical characterization of the modular
100 *Salmonella* phage endolysin Gp110. Among the many endolysins we have reported before (21-
101 23) and unpublished endolysins, Gp110 outstands in enzymatic activity (between a 1.7- and
102 364-fold increase). In addition, the catalytic domain is encoded by an un C-terminal domain of
103 unknown function (DUF3380; pfam11860). These elements prompted us to characterize Gp110
104 in more detail, including the determination of its peptidoglycan cleavage specificity.

105

106 **MATERIAL AND METHODS**

107

108 **Bacterial strains and growth conditions**

109 Wild-type *Pseudomonas aeruginosa* PAO1 strain (ATCC 15692) was kindly provided by
110 Dr. Pirnay (Lab MCT, Queen Astrid Military Hospital, Neder-Over-Heembeek, Belgium). The food
111 isolate *Salmonella enterica* serovar Typhimurium LT2 (ATCC 700720) was provided by the
112 Centre of Food and Microbial Technology of the KU Leuven (Belgium). Chemically competent
113 *Escherichia coli* TOP10 (Thermo Fischer Scientific, Waltham, MA, USA) and *E. coli*
114 BL21(DE3)pLysS (Agilent Technologies, Santa Clara, CA, USA) cells were prepared for cloning
115 and protein recombinant expression, respectively. All these strains were grown at 37°C in
116 Lysogeny Broth (LB) with shaking.

117 **Cloning, large scale expression and purification**

The sequence for a putative endolysin encoded by the *orf110* of the genome of the uncharacterized *Salmonella* phage 10 was kindly provided by Dr. K. Makhulatia, (Eliava Institute, Georgia) and has been deposited to Genbank (Accession N° KU705467). The *orf110* was amplified using Phusion High-Fidelity DNA polymerase (Thermo Fischer Scientific, Waltham, MA, USA) and primers Gp110-F (5'-ATGGCCATTCTAAAACTTGGCAACC-3') and Gp110-R (5'-GCAGAAACTCTTGTATGCTGCC-3'). The PCR product was cloned into the commercially available pEXP5-CT/TOPO® expression vector (Thermo Fischer Scientific, Waltham, MA, USA) according to the manufacturer's instructions and sequence-verified using the BigDye® Terminator v1.1 Cycle Sequencing Kit and the ABI3130 Genetic Analyzer (Life Technologies, Carlsbad, CA, USA). The pEXP5-CT/TOPO® vector provides a C-terminal 6xHis-tag for Ni-NTA purification.

Recombinant expression of Gp110 was performed in 500 ml LB at 37°C for 4 hours, using *E. coli* BL21(DE3)pLysS cells after induction during mid-exponential growth of the culture (OD_{600nm} = 0.6) with 1 mM of isopropyl-β-D-thiogalactopyranoside (IPTG). After induction, the pellet was resuspended in lysis buffer (20 mM NaH_2PO_4-NaOH, 500 mM NaCl, 50 mM imidazole, pH 7.4) and disrupted with a combination of three freeze-thawing cycles (-80°C/room temperature) and sonication (10 cycles of 30 s pulse and 30 s rest, Vibra-Cell™ Sonics and Materials, Newtown, CT, USA). Gp110 was purified using the His GraviTrap™ colum kit (GE Healthcare Life Sciences, Buckinghamshire, UK) following supplier's recommendations. Wash buffer and elution buffer were composed of 20 mM NaH_2PO_4-NaOH, 500 mM NaCl, pH 7.4 with 50 mM or 500 mM imidazole, respectively. Protein purity was estimated by SDS-PAGE. The Gp110 concentration was determined spectrophotometrically after dialyzing against

140 phosphate-buffered saline (PBS) buffer pH 7.4 using Slide-A-Lyzer® MINI dialysis units (Thermo
141 Fischer Scientific, Waltham, MA, USA). The dialyzed protein was stored at 4°C without observed
142 loss in activity.

143 **Quantification and characterization of muralytic activity**

144 The hydrolytic activity of Gp110 was quantified on *P. aeruginosa* PAO1 cells with the
145 outer membrane (OM) permeabilized by a chloroform/Tris–HCl treatment as described
146 previously (24). Briefly, mid-exponentially growing (OD_{600nm} = 0.6) *P. aeruginosa* PAO1 cells
147 were incubated in a chloroform-saturated 0.05 M Tris/HCl buffer (pH 7.7) for 45 min.
148 Afterwards, cells were washed in PBS pH 7.4 and concentrated to an OD_{600nm} of 1.5 also in PBS.
149 To determine the muralytic activity, 30 µl of Gp110 were added to 270 µl of OM permeabilized
150 *P. aeruginosa* PAO1 cells (final concentrations between 0.25 and 750 nM Gp110 for the dose-
151 dependence curve) and the resulting decrease in optical density was measured
152 spectrophotometrically (655 nm) in a Microplate Reader 680 (Bio-Rad, CA, USA). The muralytic
153 activity of Gp110 was quantified in units/µM according to a standardized method described in
154 (25).

155 The effect of pH and temperature on the lytic activity of the endolysin (final
156 concentration 2 nM) was assessed by the same method with some modifications: for the pH-
157 dependent effect, OM permeabilized *P. aeruginosa* PAO1 cells were resuspended in universal
158 pH buffer (150 mM KCl, 10 mM KH_2PO_4, 10 mM Na-citrate, 10 mM H_3BO_4) adjusted to pH values
159 between 3 and 12. To determine the effect of temperature on Gp110 activity, the endolysin
160 (final concentration 2 nM) was incubated for 10 min at either 30°C, 40°C, 50°C, 60°C, 70°C, 80°C
161 and 90°C, followed by a cooling step to room temperature. The *P. aeruginosa* PAO1 substrate

was resuspended in universal pH buffer adjusted to the optimal pH for the endolysin activity and the residual activity was tested at room temperature. For both experiments, the relative muralytic activity was calculated. All assays were performed in triplicate. Statistical analyses were performed using one-way ANOVA and the Tukey post-hoc test.

In vitro antibacterial activity

The antibacterial assay has been performed similarly as previously described (5). Mid-exponentially growing *P. aeruginosa* PAO1 and *S.* Typhimurium LT2 cells ($OD_{600\ nm}$ = 0.6) were diluted in 5 mM HEPES-NaOH (pH 7.4) to a final density of 10^6 CFU/ml. Next, 100 µl of these cultures were mixed with 50 µl of Gp110 (2.5 µM final concentration, dialyzed against phosphate-buffered saline pH 7.4) and 50 µl of 5 mM HEPES-NaOH (pH 7.4) or EDTA (0.5 mM final concentration)/malate/lactate (both with a final concentration of 10 mM) dissolved in the same buffer. As a control 100 µl of cells, 50 µl of 5 mM HEPES-NaOH (pH 7.4) and 50 µl PBS pH 7.4 were used. After incubation for 30 min at room temperature, mixtures were diluted in PBS pH 7.4 and plated on LB agar plates. The antibacterial activity is quantified after 18h incubation at 37°C as the relative inactivation in logarithmic units (= $\log_{10}(N_0/N_i)$ with N_0= number of untreated cells and N_i= number of treated cells counted after incubation).

Analysis of PG fragments solubilized by Gp110

The PG bond cleaved by Gp110 was determined using *E. coli* BW25113 Δ*lpp* PG as a substrate. PG was extracted as previously described using boiling SDS (26). A total of 500 µg of pure PG was digested overnight at 37°C with 0.5 mg/ml of Gp110 in a final volume of 200 µl. As a control, the same amount of PG was digested with 0.25 mg/ml of *Streptomyces globisporus*

mutanolysin (Sigma-Aldrich, Missouri, USA) in 25 mM phosphate buffer pH 6. After centrifugation at 20,000 × *g* for 15 min, soluble muropeptides were reduced with sodium borohydride and separated by reverse-phase HPLC (RP-HPLC) on a Hypersil aQ C18 column (3 µm; 2.1 by 200 mm; ThermoFisher Scientific, Waltham, MA, USA) coupled to an Agilent 6500 Series Q-TOF LC/MS System. Muropeptides were eluted at a flow rate of 0.25 ml/min with a 0 to 15% gradient (buffer A: 0.1% (v/v) formic acid in water; buffer B: 0.1% (v/v) formic acid in acetonitrile) applied between 6 and 40 min. *Bacillus subtilis* 168 and *Aerococcus viridans* ATCC 10400 PG were also used as substrates. Hydrolysis conditions used for *B. subtilis* and *A. viridans* PG were similar as the ones described for *E. coli* PG.

Nucleotide sequence accession number

The DNA sequence of *orf110* was deposited in GenBank under the accession number KU705467.

RESULTS

In silico analysis of Gp110

Bioinformatic analysis of the *Salmonella* phage 10 genome revealed that *orf110* encodes a 264-amino acid protein (deduced molecular mass 28.9 kDa) predicted to be a putative endolysin by HHpred (27), with some similarities (21%) with *Pseudomonas* phiKZ phage endolysin (E-value $3x10^{-20}$). Sequence similarity searches by Blastp (28) indicate that Gp110 has a 100% identity with predicted peptidoglycan binding proteins from three other *Salmonella* phages (PhiSH19, vB_SalM_SJ3 and Det7). Conserved domains analysis by Pfam (29) showed

205 that Gp110 has a modular structure, with an N-terminal PG_binding_1 domain (pfam01471)
206 and a C-terminal DUF3380 domain (pfam11860). PG_binding_1 from Gp110 has specific
207 repeated motifs (DGIFGKAT and DGIAGPKT), a feature that seems to be common in proteins
208 interacting with repetitive structures like peptidoglycan (30). These motifs match the consensus
209 sequence (D-G-(Pho)$_2$-G-K/N-G/N-T; Pho = hydrophobic amino acid) previously found in other
210 endolysins with a Gram-negative background (21, 23). The DUF3380 domain is found in viruses
211 and bacteria, normally associated to PG_binding_1, and belongs to a family of functionally
212 uncharacterized proteins.

213 **Biochemical characterization of Gp110 peptidoglycan-degrading activity**

214 The predicted PG hydrolytic activity of Gp110 was confirmed and characterized using
215 OM-permeabilized *P. aeruginosa* cells as a substrate (24) to allow the endolysin to reach the PG
216 layer and exert its enzymatic activity, which is measured by a turbidity assay (25).
217 Incubation of Gp110 with the substrate in universal buffer adjusted to different pHs showed
218 that this enzyme is active at pH values ranging from 4 to 11, maintaining between 16.0 ± 2.67%
219 and 47.0 ± 6.9% of its maximal activity at pH 4 and 11, respectively. The highest activity is
220 achieved in the pH range of 6-9 (between 83.6 ± 13.0% and 93.6 ± 1.4%, respectively), (Figure
221 1.). Of note is the significant decrease in activity below pH 6 and above pH 9, losing >50% of the
222 activity ($p < 0.05$). Gp110 activity shows no significant loss in activity upon a 10 min heat
223 exposure to temperatures between 20 and 60°C. Above 60°C the activity gradually decreases
224 with a remaining activity of 26.7 ± 7.0% after 10 min at 90°C (Figure 2.).

The specific activity of Gp110 was calculated at optimal pH under substrate-saturating conditions from the slope of the linear regression of the corresponding dose-dependent saturation curve as previously described (25) (Figure 3.). A linear dose-response was observed between 0.25 and 1.25 nM Gp110. According to this method the specific activity of Gp110 was calculated to be 34,240 units/µM, corresponding to a complete clarification of a turbid cell culture in approximately 10 and 20 minutes with 10 and 1 nM Gp110, respectively.

Gp110 antibacterial activity

The *in vitro* antibacterial activity of the endolysin was tested against *P. aeruginosa* PAO1 and *S.* Typhimurium LT2 in the presence or absence of EDTA, malate and lactate as OM permeabilizers (31). As expected, the activity of 2.5 µM Gp110 alone was insignificant against both strains. However, addition of 0.5 mM EDTA, together with the 2.5 µM Gp110, resulted in a reduction of 2.74 ± 0.11 log-units in case of *P. aeruginosa* PAO1, corresponding to nearly a 99.9% reduction in the number of viable cells, and 0.38 ± 0.18 log units in case of *S.* Typhimurium LT2 cells. The addition of EDTA alone reduced the cell number with 0.62 ± 0.06 and 0.28 ± 0.29 log units for *P. aeruginosa* and *S.* Typhimurium, respectively. Other OM permeabilizers such us malate and lactate, did not improve significantly the antibacterial activity of the endolysin against *S.* Typhimurium LT2 (data not shown).

Determination of Gp110 PG cleavage specificity by LC-MS

To determine the PG bond cleaved by the DUF3380 domain in Gp110, purified *E. coli* PG was incubated in the presence of recombinant Gp110 and soluble fragments were analyzed by rp-HPLC coupled to MS (Figure 4.). The muropeptide profile obtained for Gp110 is very similar

246 to the one obtained for the *N*-acetylmuramidase mutanolysin (Figure 4A.), suggesting that
247 Gp110 cleaves the β-1,4 glycosidic bonds of the PG sugar backbone. The major monomer (peak
248 4, Figure 4A.) generated after the digestion of the PG with Gp110 yielded a peak with an *m/z* at
249 942.415 matching the theoretical value expected for a disaccharide-tetrapeptide (Figure 4B.)
250 thus confirming the glycosidase activity. To determine whether DUF3380 possesses *N*-
251 acetylmuramidase or *N*-acetylglucosaminidase activity, tandem mass spectrometry was further
252 performed on the major monomer (peak 4). The fragmentation event leading to the loss of a
253 non-reduced GlcNAc residue (203.078 atomic mass units) indicated that Gp110 displays *N*-
254 acetylmuramidase activity (Figure 4C.) instead of an *N*-acetylglucosaminidase activity which
255 would involve the loss of 223.106 atomic mass units corresponding to a reduced GlcNAc (32).
256 Interestingly, Gp110 also displays enzymatic activity against *B. subtilis* and *A. viridans* PG
257 (Figure S1.). This indicates that this enzyme can cleave PG containing amidated *meso*-DAP or L-
258 lysine residues at position 3 of the peptide stems. Unlike *E. coli* PG digestion products, which
259 essentially contains tetrapeptides stems, both *B. subtilis* and *A. viridans* digestion products
260 mostly contain tripeptide stems.
261
262 **DISCUSSION**
263
264 In this study, we describe a new endolysin from an uncharacterized phage infecting the
265 Gram-negative pathogen *S.* Typhimurium. *In silico* analysis indicates that this endolysin has a
266 modular structure harbouring a DUF3380 EAD at the C-terminus and a PG_binding_1 CBD at the
267 N-terminus. This modular structure is a common feature in endolysins with a Gram-positive

268 background (7) but remains rare in endolysins with a Gram-negative background, which are
269 mostly globular (22). Interestingly, the PG_binding_1 is also almost restricted to Gram-positive
270 related endolysins (8), although with a few exceptions (*Salmonella* phage PVP-SE1;
271 *Pseudomonas* phages phiKZ, EL, 201phi21, and OBP; and phages infecting *Burkholderia* and
272 *Erwinia*) (21, 23, 33). The DUF3380 domain has been only predicted in endolysins from phages
273 infecting *Burkholderia*, *Pseudomonas* and *Erwinia* (33) and it is classified as a pfam domain of
274 unknown function. Noteworthy is the fact that there is another domain with unknown function
275 (DUF3597, pfam12200) associated to an endolysin but it has only been detected in the
276 endolysin from the *Listeria* phage A118 (8, 33). According to our results, the biochemical
277 analysis of DUF3380 using pure PG revealed that this domain has *N*-acetylmuramidase (EC
278 3.2.1.17) activity, cleaving the β-(1,4) bonds between *N*-acetylmuramic acid and *N*-
279 acetylglucosamine in the sugar backbone of the PG. In addition, Gp110 can cleave PG with
280 distinct PG compositions such as in *B. subtilis* and *A. viridans*. Up to 690 proteins have a
281 predicted DUF3380 domain, among which the large majority is present in bacterial proteins and
282 only 10% are encoded by bacteriophages (InterPro database), which may indicate a horizontal
283 transfer. According to the Pfam database, this domain is commonly found associated with a PG
284 binding domain, however, it is also found in one-domain proteins (InterPro). Recently, using
285 remote homology detection methods, such as FFAS03 (http://ffas.sanfordburnham.org) and
286 data from publications collected by PubServer (http://pubserver.burnham.org), DUF3380 was
287 designated as a family of cell-wall lytic enzymes (34) which is in accordance with our results and
288 we have confirmed it by cleavage sites determination in the PG. Alignments of some sequences
289 containing this DUF3380 domain show that Gp110 E101 residue is conserved, suggesting this is

290 the catalytic residue. Moreover, this catalytic residue is followed by serine which is a common
291 feature in lysozymes. However, three-dimensional structure predictions of the DUF3380
292 domain show a low homology with the tertiary structure of other lysozymes which suggests a
293 completely different fold. On the contrary, the CBD is strongly conserved, showing high
294 homology with PG binding domains of other hydrolases.

295 Comparing the Gp110 specific activity with other endolysins analyzed using the same
296 method under optimal conditions (Table 1.), this endolysin shows the highest enzymatic activity
297 described to date, 1.7-fold more active than the second most active endolysin, OBPgp279,
298 which has also a modular structure and a predicted lysozyme-like muramidase activity (21).
299 Compared to other *Salmonella* phage endolysins, Gp110 has a 2.5- and 24.8-fold higher activity
300 than PVP-SE1gp146 and PsP3gp10, respectively, and even a 85.6-fold higher activity than Lys68
301 (Table 1.). The optimal pH for Gp110 activity is pH 8 (Figure 1.), in contrast to other previously
302 described endolysins with Gram-negative background, which have an optimal activity at neutral
303 pH (21, 23, 22, 35). It is noteworthy that three other endolysins from phages infecting *S.*
304 Typhimurium have been reported to show an optimal activity at pH 9.5 (phage SPN1S
305 endolysin) and pH 8.5 (phage SPN9CC endolysin and Lys394) (36-38), whereas endolysins from
306 phages infecting *S.* Enteritidis (PVP-SE1gp146, Lys68 and PsP3gp10) have an optimal activity at
307 pH 7 (21, 22, 35). In terms of temperature resistance, Gp110 remains active after treatment at
308 temperatures across the tested range (20-90°C), showing no significant activity loss up to 10
309 min heat treatments at 60°C and only gradually decreasing at higher temperatures (Figure 2.).
310 Whereas most endolysins irreversibly lose their enzymatic activity upon exposure to
311 temperatures around 50°C (37, 39-43), some thermoresistant endolysins have also been

312 described, including Lys68 from *Salmonella* phage phi68 (35), gp146 from *Salmonella* phage
313 PVP-SE1 (21), the endolysins from bacteriophages Ph2119 and vB_Tsc2631 infecting the
314 thermophile *Thermus scotoductus* (44, 45) and the lysin from deep-sea thermophilic
315 bacteriophage GVE2 (46). Gp110 shows an intermediate profile with an elevated temperature
316 resistance compared to most mesophilic endolysins, which may correlate to an increased shelf
317 life of Gp110.

318 To verify if the biochemical activity of DUF3380 is translated into an antibacterial
319 activity, Gp110 (2.5 µM) was tested against *P. aeruginosa* PAO1 and *S.* Typhimurium LT2 cells in
320 the presence or absence of 0.5 mM EDTA as OM permeabilizer. As expected, the antibacterial
321 activity of the protein without EDTA was very low against both strains due to the OM protective
322 effect. However, in the presence of EDTA the number of *P. aeruginosa* PAO1 viable cells was
323 reduced in a 99.9% approximately. The synergistic effect of EDTA in the antibacterial activity of
324 endolysins with a Gram-negative background was first described with the endolysin EL188 and
325 *P. aeruginosa* (47). The mild antibacterial activity of the endolysin/EDTA combination against *S.*
326 Typhimurium LT2 cells was also observed for other endolysins (21, 36, 38) and can be explained
327 by the lower degree of phosphorylation in *Salmonella* lipopolysaccharide (LPS) molecules
328 compared to *Pseudomonas* LPS. As a consequence, *S.* Typhimurium LT2 has a significantly lower
329 amount of stabilizing divalent cations, resulting in a lower susceptibility to the EDTA
330 permeabilization, compared to *P. aeruginosa*.

331 Overall, we have characterized an endolysin with the highest enzymatic activity against
332 Gram-negative peptidoglycan reported to date. This high Gp110 activity can be explained by its
333 enzymatically active domain, DUF3380, which was biochemically demonstrated to have *N*-

acetylmuramidase activity and shows a low degree of homology with lysozymes. These features render Gp110 a novel attractive candidate for engineering to provide the enzyme with outer membrane permeabilizing and consequently antibacterial properties.

337 **REFERENCES**

338 1. **Young R.** 1992. Bacteriophage lysis: mechanism and regulation. Microbiol Rev **56:**430-
339 481.

340 2. **Roach DR, Donovan DM.** 2015. Antimicrobial bacteriophage-derived proteins and
341 therapeutic applications. Bacteriophage **5:**e1062590.

342 3. **Briers Y, Walmagh M, Van Puyenbroeck V, Cornelissen A, Cenens W, Aertsen A,**
343 **Oliveira H, Azeredo J, Verween G, Pirnay JP, Miller S, Volckaert G, Lavigne R.** 2014.
344 Engineered endolysin-based "Artilysins" to combat multidrug-resistant gram-negative
345 pathogens. MBio **5:**e01379-01314.

346 4. **Briers Y, Lavigne R.** 2015. Breaking barriers: expansion of the use of endolysins as novel
347 antibacterials against Gram-negative bacteria. Future Microbiol **10:**377-390.

348 5. **Briers Y, Walmagh M, Grymonprez B, Biebl M, Pirnay JP, Defraine V, Michiels J, Cenens**
349 **W, Aertsen A, Miller S, Lavigne R.** 2014. Art-175 is a highly efficient antibacterial against
350 multidrug-resistant strains and persisters of *Pseudomonas aeruginosa*. Antimicrob
351 Agents Chemother **58:**3774-3784.

352 6. **Rodríguez-Rubio L, Gutiérrez D, Donovan DM, Martínez B, Rodríguez A, García P.** 2015.
353 Phage lytic proteins: biotechnological applications beyond clinical antimicrobials. Crit
354 Rev Biotechnol **10:**1-11.

355 7. **Nelson DC, Schmelcher M, Rodríguez-Rubio L, Klumpp J, Pritchard DG, Dong S,**
356 **Donovan DM.** 2012. Endolysins as antimicrobials. Adv Virus Res **83:**299-365.

8. **Oliveira H, Melo LD, Santos SB, Nobrega FL, Ferreira EC, Cerca N, Azeredo J, Kluskens LD.** 2013. Molecular aspects and comparative genomics of bacteriophage endolysins. J Virol **87:**4558-4570.

9. **Loessner MJ, Kramer K, Ebel F, Scherer S.** 2002. C-terminal domains of *Listeria monocytogenes* bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. Mol Microbiol **44:**335-349.

10. **Schleifer KH, Kandler O.** 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol Rev **36:**407-477.

11. **Vollmer W, Blanot D, de Pedro MA.** 2008. Peptidoglycan structure and architecture. FEMS Microbiol Rev **32:**149-167.

12. **Navarre WW, Ton-That H, Faull KF, Schneewind O.** 1999. Multiple enzymatic activities of the murein hydrolase from staphylococcal phage phi11. Identification of a D-alanyl-glycine endopeptidase activity. J Biol Chem **274:**15847-15856.

13. **Paradis-Bleau C, Cloutier I, Lemieux L, Sanschagrin F, Laroche J, Auger M, Garnier A, Levesque RC.** 2007. Peptidoglycan lytic activity of the *Pseudomonas aeruginosa* phage phiKZ gp144 lytic transglycosylase. FEMS Microbiol Lett **266:**201-209.

14. **Pritchard DG, Dong S, Baker JR, Engler JA.** 2004. The bifunctional peptidoglycan lysin of *Streptococcus agalactiae* bacteriophage B30. Microbiology **150:**2079-2087.

15. **Becker SC, Dong S, Baker JR, Foster-Frey J, Pritchard DG, Donovan DM.** 2009. LysK CHAP endopeptidase domain is required for lysis of live staphylococcal cells. FEMS Microbiol Lett **294:**52-60.

16. **Loessner MJ, Gaeng S, Wendlinger G, Maier SK, Scherer S.** 1998. The two-component lysis system of *Staphylococcus aureus* bacteriophage Twort: a large TTG-start holin and an associated amidase endolysin. FEMS Microbiol Lett **162:**265-274.

17. **Fischetti VA, Zabriskie JB, Gotschlich EC.** Physical, chemical, and biological properties of type 6 M-protein extracted with purified streptococcal phage-associated lysin, p 26. *In* Haverkorn MJ (ed), Streptococcal disease and the community. Excerpta Medica, Amsterdam.

18. **Inouye M, Arnheim N, Sternglanz R.** 1973. Bacteriophage T7 lysozyme is an N-acetylmuramyl-L-alanine amidase. J Biol Chem **248:**7247-7252.

19. **Bienkowska-Szewczyk K, Lipinska B, Taylor A.** 1981. The R gene product of bacteriophage lambda is the murein transglycosylase. Mol Gen Genet **184:**111-114.

20. **Pritchard DG, Dong S, Kirk MC, Cartee RT, Baker JR.** 2007. LambdaSa1 and LambdaSa2 prophage lysins of *Streptococcus agalactiae*. Appl Environ Microbiol **73:**7150-7154.

21. **Walmagh M, Briers Y, dos Santos SB, Azeredo J, Lavigne R.** 2012. Characterization of modular bacteriophage endolysins from Myoviridae phages OBP, 201phi2-1 and PVP-SE1. PLoS One **7:**e36991.

22. **Walmagh M, Boczkowska B, Grymonprez B, Briers Y, Drulis-Kawa Z, Lavigne R.** 2013. Characterization of five novel endolysins from Gram-negative infecting bacteriophages. Appl Microbiol Biotechnol **97:**4369-4375.

23. **Briers Y, Volckaert G, Cornelissen A, Lagaert S, Michiels CW, Hertveldt K, Lavigne R.** 2007. Muralytic activity and modular structure of the endolysins of *Pseudomonas aeruginosa* bacteriophages phiKZ and EL. Mol Microbiol **65:**1334-1344.

24. **Lavigne R, Briers Y, Hertveldt K, Robben J, Volckaert G.** 2004. Identification and characterization of a highly thermostable bacteriophage lysozyme. Cell Mol Life Sci **61:**2753-2759.

25. **Briers Y, Lavigne R, Volckaert G, Hertveldt K.** 2007. A standardized approach for accurate quantification of murein hydrolase activity in high-throughput assays. J Biochem Biophys Methods **70:**531-533.

26. **Glauner B.** 1988. Separation and quantification of muropeptides with high-performance liquid chromatography. Anal Biochem **172:**451-464.

27. **Soding J, Biegert A, Lupas AN.** 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res **33:**W244-248.

28. **Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:**3389-3402.

29. **Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A.** 2010. The Pfam protein families database. Nucleic Acids Res **38:**D211-222.

30. **Wren BW.** 1991. A family of clostridial and streptococcal ligand-binding proteins with conserved C-terminal repeat sequences. Mol Microbiol **5:**797-803.

31. **Vaara M.** 1992. Agents that increase the permeability of the outer membrane. Microbiol Rev **56:**395-411.

421 32. **Eckert C, Lecerf M, Dubost L, Arthur M, Mesnage S.** 2006. Functional analysis of AtlA,
422 the major *N*-acetylglucosaminidase of *Enterococcus faecalis*. J Bacteriol **188:**8513-8519.
423 33. **Vidová B, Šramková Z, Tišáková L, Oravkinová M, Godány A.** 2014. Bioinformatics
424 analysis of bacteriophage and prophage endolysin domains. Biologia **69:**541-556.
425 34. **Jaroszewski L, Koska L, Sedova M, Godzik A.** 2014. PubServer: literature searches by
426 homology. Nucleic Acids Res **42:**W430-435.
427 35. **Oliveira H, Thiagarajan V, Walmagh M, Sillankorva S, Lavigne R, Neves-Petersen MT,**
428 **Kluskens LD, Azeredo J.** 2014. A thermostable *Salmonella* phage endolysin, Lys68, with
429 broad bactericidal properties against gram-negative pathogens in presence of weak
430 acids. PLoS One **9:**e108376.
431 36. **Lim JA, Shin H, Kang DH, Ryu S.** 2012. Characterization of endolysin from a *Salmonella*
432 Typhimurium-infecting bacteriophage SPN1S. Res Microbiol **163:**233-241.
433 37. **Legotsky SA, Vlasova KY, Priyma AD, Shneider MM, Pugachev VG, Totmenina OD,**
434 **Kabanov AV, Miroshnikov KA, Klyachko NL.** 2014. Peptidoglycan degrading activity of
435 the broad-range *Salmonella* bacteriophage S-394 recombinant endolysin. Biochimie **107**
436 **Pt B:**293-299.
437 38. **Lim JA, Shin H, Heu S, Ryu S.** 2014. Exogenous lytic activity of SPN9CC endolysin against
438 gram-negative bacteria. J Microbiol Biotechnol **24:**803-811.
439 39. **Filatova LY, Becker SC, Donovan DM, Gladilin AK, Klyachko NL.** 2010. LysK, the enzyme
440 lysing *Staphylococcus aureus* cells: specific kinetic features and approaches towards
441 stabilization. Biochimie **92:**507-513.

40. **Varea J, Monterroso B, Saiz JL, Lopez-Zumel C, Garcia JL, Laynez J, Garcia P, Menendez M.** 2004. Structural and thermodynamic characterization of Pal, a phage natural chimeric lysin active against pneumococci. J Biol Chem **279:**43697-43707.

41. **Sanz JM, Garcia JL, Laynez J, Usobiaga P, Menendez M.** 1993. Thermal stability and cooperative domains of CPL1 lysozyme and its NH2- and COOH-terminal modules. Dependence on choline binding. J Biol Chem **268:**6125-6130.

42. **Obeso JM, Martinez B, Rodriguez A, Garcia P.** 2008. Lytic activity of the recombinant staphylococcal bacteriophage PhiH5 endolysin active against *Staphylococcus aureus* in milk. Int J Food Microbiol **128:**212-218.

43. **Heselpoth RD, Nelson DC.** 2012. A new screening method for the directed evolution of thermostable bacteriolytic enzymes. J Vis Exp doi:10.3791/4216.

44. **Plotka M, Kaczorowska AK, Stefanska A, Morzywolek A, Fridjonsson OH, Dunin-Horkawicz S, Kozlowski L, Hreggvidsson GO, Kristjansson JK, Dabrowski S, Bujnicki JM, Kaczorowski T.** 2014. Novel highly thermostable endolysin from *Thermus scotoductus* MAT2119 bacteriophage Ph2119 with amino acid sequence similarity to eukaryotic peptidoglycan recognition proteins. Appl Environ Microbiol **80:**886-895.

45. **Plotka M, Kaczorowska AK, Morzywolek A, Makowska J, Kozlowski LP, Thorisdottir A, Skirnisdottir S, Hjorleifsdottir S, Fridjonsson OH, Hreggvidsson GO, Kristjansson JK, Dabrowski S, Bujnicki JM, Kaczorowski T.** 2015. Biochemical Characterization and Validation of a Catalytic Site of a Highly Thermostable Ts2631 Endolysin from the *Thermus scotoductus* Phage vB_Tsc2631. PLoS One **10:**e0137374.

46. **Ye T, Zhang X.** 2008. Characterization of a lysin from deep-sea thermophilic bacteriophage GVE2. Appl Microbiol Biotechnol **78:**635-641.

47. **Briers Y, Walmagh M, Lavigne R.** 2011. Use of bacteriophage endolysin EL188 and outer membrane permeabilizers against *Pseudomonas aeruginosa*. J Appl Microbiol **110:**778-785.

471 **TABLES**

472

473 **Table 1**. Comparison of Gp110 specific activity with other endolysins against *P. aeruginosa*

474 PAO1. All specific activities were calculated following the method described in (22).

Endolysin	Activity (units/µM)	Structure	Reference
Gp110	34240	Modular	This work
OBPgp279	19979	Modular	(29)
LysEC8	17103	Globular	Unpublished results
PVP-SE1gp146	13614	Modular	(29)
EL188	4735	Modular	(28)
201φ2-1gp229	4469	Modular	(29)
KZ144	2058	Modular	(28)
PsP3gp10	1380	Globular	(33)
P2gp09	829	Globular	(33)
BcepC6Bgp22	786	Globular	(33)
Lys68	400	Globular	(36)
CR8gp3.5	315	Globular	(3)
K11gp3.5	134	Globular	(33)
KP32gp15	117	Globular	(33)
LysAci7	94	Globular	Unpublished results

475

476

477 **FIGURES**

478

479 **Figure 1**. pH dependence of Gp110 enzymatic activity. The muralytic activity of Gp110 is shown

480 against OM permeabilized *P. aeruginosa* PAO1 cells resuspended in a universal buffer adjusted

481 to different pH values. Activity is expressed relative to the maximal muralytic activity at the

482 optimal pH (pH 8). Each bar represents the mean of triplicate experiments, and error bars

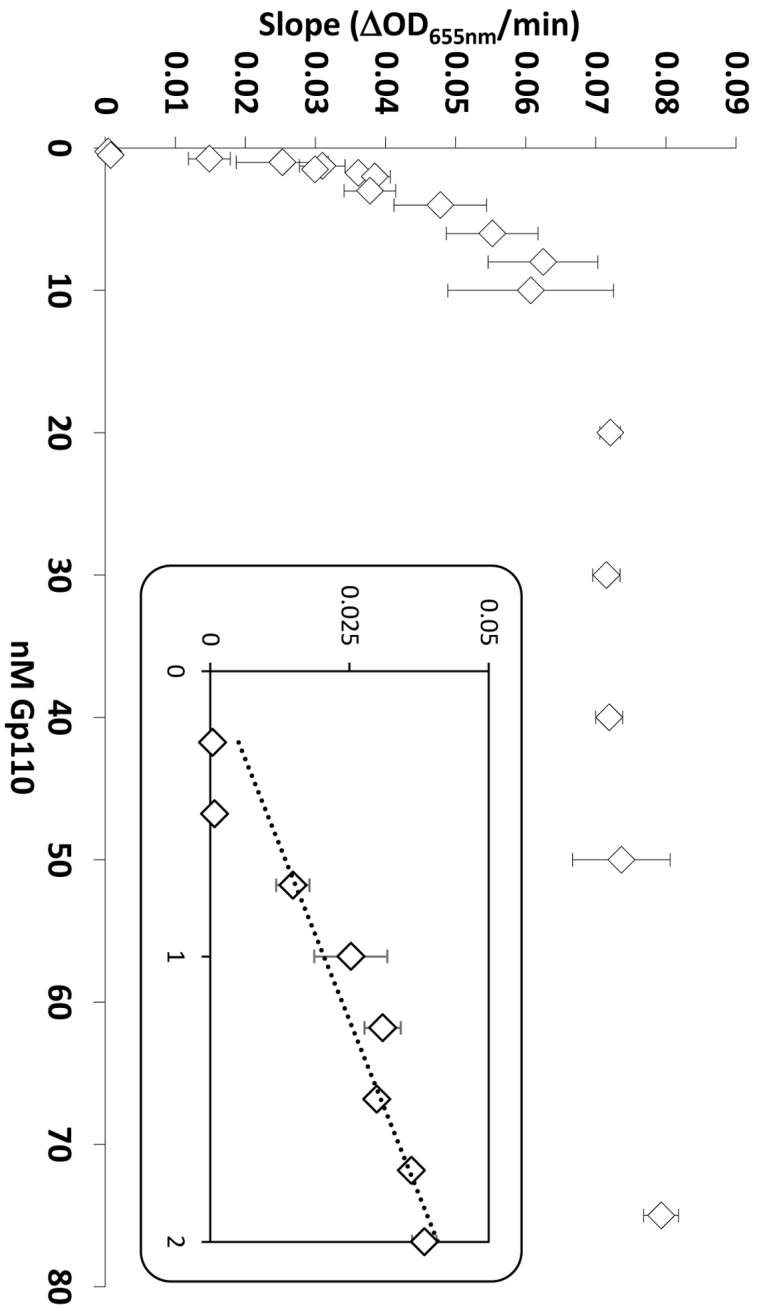
483 indicate the standard deviation. A one-way ANOVA and Tukey post-hoc test indicated that
484 there were no statistically significant differences between the activities at the pH range of 6-9
485 (all $p > .05$).

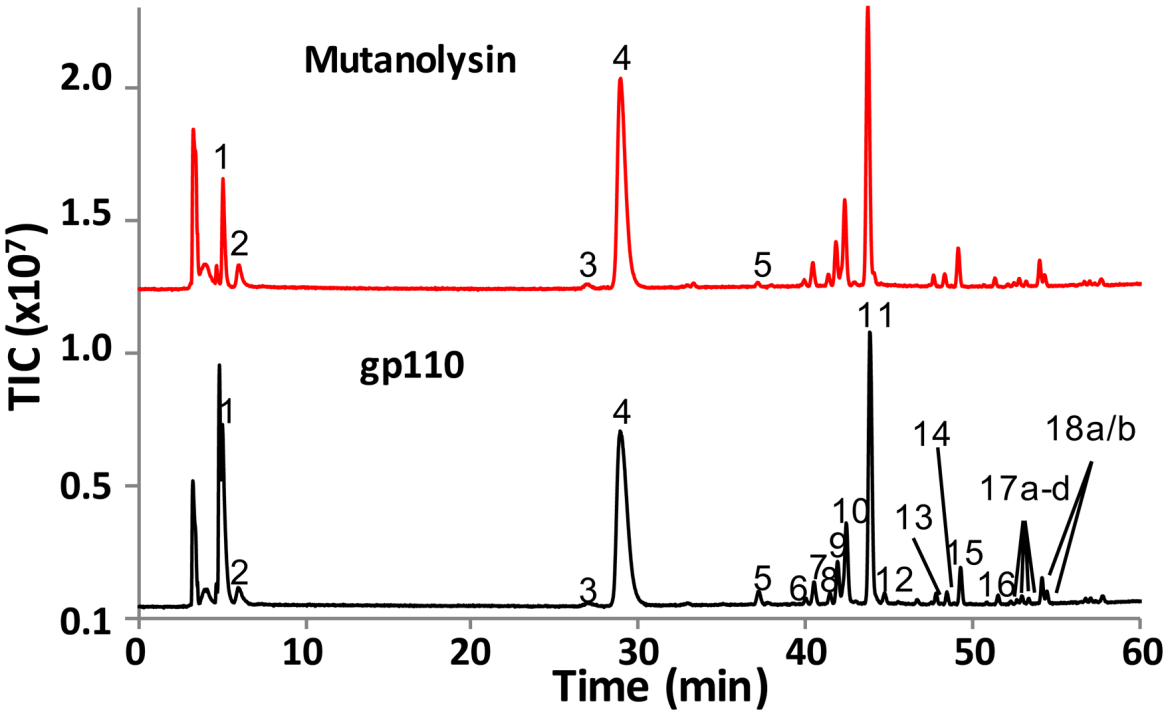
486 **Figure 2.** Temperature resistance of Gp110. The residual enzymatic activity of Gp110 was
487 analyzed after 10 min incubation of the endolysin at different temperatures, ranging from 20°C
488 to 90°C. PG hydrolase activity is expressed relative to the highest measured activity.

489 **Figure 3.** Saturation curve of Gp110 at optimal pH (pH 8). The muralytic activity was quantified
490 against OM permeabilized *P. aeruginosa* PAO1 cells according to Briers *et al*. (2007). The X- and
491 Y-axes display the amount of Gp110 (in nM) added and the corresponding activity (OD_{655}/min)
492 measured, respectively. Each data point shows the average and error bars of three replicates.
493 The insert zooms in on the region with a linear relationship between activity and enzyme
494 concentration (substrate saturating conditions)

495 **Figure 4.** Determination of Gp110 cleavage specificity. (A) LC-MS analysis of *E. coli* BW25113
496 D*lpp* peptidoglycan digested by mutanolysin and recombinant Gp110. Soluble muropeptides
497 were reduced and analysed by rp-HPLC coupled to MS. Peaks corresponding to *m/z* values
498 matching previously identified muropeptides are numbered. (B) Inferred structures, theoretical
499 monoisotopic masses, theoretical and observed *m/z* values of peaks identified in (A). (C) LC-
500 MS/MS analysis of the major disaccharide-peptide (peak 4) solubilised by Gp110. The
501 fragmentation pattern of the $[M+H]^+$ ion at *m/z* 942.414 was typical of a disaccharide-
502 tetrapeptide (DS-Tetra). The fragmentation event leading to the loss of a nonreduced GlcNAc
503 residue (203.078) indicates that Gp110 displays *N*-acetylmuramidase (lysozyme) activity. The

sequence of peptide fragments is indicated above their respective *m/z* values (*boxed*). *A*, L-Ala or D-Ala; *a*, C-terminal D-Ala; *m-DAP*, *meso*-diaminopimelic acid; *E*, g-D-Glu; M^R, reduced MurNAc; *G*, GlcNAc. TIC, Total Ion Count.





Peak	Inferred structure	Monoisotopic Mass (Da)	*m/z*		
			Calculated	Observed for gp110	Observed for mutanolysin
1	GM^R-Tri	870.371	871.378	871.378	871.378
2	GM^R-Tri-Gly	927.392	928.400	928.398	928.398
3	GM^R-Di	698.286	699.294	699.292	699.292
4	GM^R-Tetra	941.408	942.415	942.415	942.415
5	GM^R-Tetra-lactyl-Tetra	1384.609	693.312[a]	693.310	693.309
6	GM^R-Tri-Gly-GM^R-Tri	1779.752	890.884[a]	1779.745	890.880
7	GM^R-Tri-GM^R-Tri	1722.731	1723.738	1723.732	1723.736
8	GM^R-Tri-Gly-GM^R-Tetra	1850.789	926.402[a]	1850.784	926.400
9	GM^R-Tetra-GM^R-Tri	1793.768	1794.775	1794.771	1794.770
10	GM^R-Tetra-GM^R-Tri	1793.768	1794.775	1794.771	1794.771
11	GM^R-Tetra-GM^R-Tetra	1864.805	1865.813	1865.809	1865.810
12	GM^R-Tri-Tetra	1456.630	1457.638	1457.635	ND
13	GM^A-Tetra	921.381	922.389	922.387	922.388
14	GM^R-Tetra-GM^R-Tetra-GM^R-Tri	2717.165	906.729[b]	906.726[b]	907.726[b]
15	GM^R-Tetra-GM^R-Tetra-GM^R-Tetra	2788.202	930.408[b]	930.406[b]	930.406[b]
16	GM^R-Tri-GM^R-Tri	1702.704	852.360[a]	852.357[a]	852.357[a]
17a-d	GM^A-Tri-GM^R-Tetra	1773.742	887.879[a]	887.875[a]	887.875[a]
				887.876[a]	887.875[a]
				887.876[a]	887.876[a]
				887.875[a]	887.876[a]
18a-b	GM^R-Tetra- GM^A-Tetra	1844.779	1845.786	1845.780	1845.781
			923.397[a]	923.394[a]	923.394[a]

[a], $[M+2H]^{2+}$ adduct; [b], $[M+3H]^{3+}$ adduct